Short-term Efficacy of Upper-Extremity Exercise Training in Patients With Chronic Airway Obstruction: A Systematic Review

Stefania Costi, Mauro Di Bari, Paolo Pillastrini, Roberto D'Amico, Ernesto Crisafulli, Cinzia Arletti, Leonardo M Fabbri, Enrico M Clini

Background, Objectives, and Measurements. Patients with chronic airway obstruction (CAO) frequently experience dyspnea and fatigue during activities performed by accessory muscles of ventilation, which competitively participate in arm elevation. This systematic review of randomized controlled trials (RCTs) concerning patients with CAO addresses the effects of upper-extremity exercise training (UEET), added to lower-extremity training or comprehensive pulmonary rehabilitation, on the following patient-centered outcomes: exercise capacity, symptoms, ability to perform daily activities, and health-related quality of life.

Methods. Studies were retrieved using comprehensive database and hand-search strategies. Two independent reviewers determined study eligibility based on inclusion criteria. A detailed description of treatments was mandatory. Reviewers rated study quality and extracted information on study methods, design, intervention, and results.

Results. Forty publications were evaluated. Four RCTs met the inclusion criteria but had serious methodological limitations, which introduce possible biases that reduce their internal validity. The outcomes measured were heterogeneous, and the results were inconsistent regarding maximal exercise capacity, dyspnea, and health-related quality of life. No effect of UEET was demonstrated for measures of arm fatigue.

Limitations and Conclusions. The limited methodological quality of the studies retrieved prevented us from performing a meta-analysis, the results of which could be misleading. This systematic review shows that there is limited evidence examining UEET and that the evidence available is of poor quality. Therefore, a recommendation for the inclusion or exclusion of UEET in pulmonary rehabilitation programs for individuals with CAO is not possible. Further research is needed to definitively ascertain the effects of this training modality on patient-centered outcomes.

S Costi, PT, MSc, is Physical Therapist, Section of Respiratory Disease, Department of Oncology, Haematology, and Respiratory Disease, University of Modena and Reggio Emilia, Via del Pozzo 74, 41100 Modena, Italy. Address all correspondence to Ms Costi at: stefania.costi@unimore.it.

M Di Bari, MD, is Doctor, Department of Critical Care Medicine and Surgery, Unit of Gerontology and Geriatrics, University of Florence and Azienda Ospedaliero-Universitaria Careggi, Florence, Italy.

P Pillastrini, PT, MSc, is Physical Therapist, Department of Neurological Sciences, Occupational Medicine Unit, S Orsola-Malpighi Hospital, University of Bologna, Bologna, Italy.

R D'Amico, PhD, is Statistician, Section of Respiratory Disease, Department of Oncology, Haematology, and Respiratory Disease, University of Modena and Reggio Emilia.

E Crisafulli, MD, is Doctor, Department of Pulmonary Rehabilitation, Ospedale Villa Pineta, Pavullo, Modena, Italy.

C Arletti, PT, BSc, Physical Therapist, Anni Azzurri, Ducale 1, 2, 3, Modena, Italy.

LM Fabbri, MD, is Doctor, Section of Respiratory Disease, Department of Oncology, Haematology, and Respiratory Disease, University of Modena and Reggio Emilia.

EM Clini, MD, FCCP, is Doctor, Section of Respiratory Disease, Department of Oncology, Haematology, and Respiratory Disease, University of Modena and Reggio Emilia, and Department of Pulmonary Rehabilitation, Ospedale Villa Pineta.

[Costi S, Di Bari M, Pillastrini P, et al. Short-term efficacy of upper-extremity exercise training in patients with chronic airway obstruction: a systematic review. *Phys Ther.* 2009;89:443–455.]

The upper extremities (UEs) play an important role in performing many activities of daily living (ADL), both in the domain of basic self-care and in everyday jobs. Patients with chronic airway obstruction (CAO) frequently experience marked dyspnea and fatigue when performing these tasks,[1,2] which commonly require unsupported arm work and, therefore, pose a unique challenge to these individuals, whose upper-limb muscles are frequently recruited as accessory inspiratory muscles.[3–5] During unsupported arm exercise, the participation of these muscles in ventilation decreases, and there is a shift of respiratory work to the diaphragm, which is commonly weakened and has a reduced functional capacity in these patients.[6] This shift is associated with thoracic-abdominal desynchronization, severe dyspnea, and premature termination of exercise.[1,7,8]

The effectiveness of pulmonary rehabilitation (PR) programs has been well documented in patients with CAO, with consistent and clinically significant improvements in exercise capacity, symptoms, and health-related quality of life (HRQoL).[9] However, such programs primarily focus on lower-extremity (LE) exercise training.[10,11] Because training effects are specific to the limb trained, it seems reasonable to assume—but it so far remains unproven—that, in patients with CAO, upper-extremity exercise training (UEET) may improve functional status and reduce symptoms while performing ADL.

Recent guidelines from the American College of Chest Physicians[11] recommend the introduction of unsupported endurance training of the UEs in PR programs. We think that this recommendation relies on limited evidence available from both randomized[12–17] and nonrandomized[18–20] studies conducted over recent decades. However, to our knowledge, no systematic review has ever been conducted on this topic.

We undertook this systematic review of randomized controlled trials (RCTs) to clarify the effect of UEET, implemented over and above standard treatment or lower-extremity exercise training (LEET), on patient-centered outcomes, such as exercise capacity, symptoms, ability to perform ADL, and HRQoL in patients with CAO.

Method

Data Sources and Searches

We performed a computer-based search, querying Ovid MEDLINE (1950 to March 2007), CINAHL (Cumulative Index to Nursing and Allied Health, 1982 to March 2007), EMBASE (1980 to March 2007), PEDro (Physiotherapy Evidence Database), and the Cochrane Central Register of Controlled Trials for original research articles published in English, Italian, and Spanish.

Search terms and strategies were as follows: (chronic airway obstruction OR pulmonary diseases chronic obstructive) AND (exercise therapy OR exercise OR rehabilitation OR physical therapy OR physiotherapy OR training) AND (arm OR upper extremities).

In addition, reference lists of relevant research articles were reviewed for pertinent studies. Abstracts presented at international meetings (American Thoracic Society, 2001–2007, European Respiratory Society, 2001–2007) also were hand-searched, and the authors of appropriate abstracts were contacted to obtain from them complete, unpublished data. Finally, PR experts were contacted to locate any further, unpublished material.

Study Selection

The following criteria were used to select trials for inclusion in the review.

Design. We considered for inclusion RCTs only.

Target population. Trials were considered when they enrolled patients with a diagnosis of moderate, severe, or very severe CAO. The criteria used for this purpose were the best recorded ratio of forced expiratory volume in 1 second (FEV_1) to forced vital capacity (FVC) of less than 0.7, associated with the best recorded FEV_1 of less than 80% of the predicted value,[21] and a clinical diagnosis of CAO.

Intervention. As recommended by the major scientific societies in this field,[10] we selected any inpatient, outpatient, or home-based PR programs that included at least 20 sessions for a minimum frequency of 3 times a week. Both supervised and unsupervised home sessions were considered acceptable. The program had to include supported or unsupported UEET as the experimental intervention. A detailed description of the experimental intervention was mandatory.

Control. Randomized controlled trials were included only when they compared UEET with treatments not specifically aimed at improving UE exercise capacity. The control group could receive standard training consisting of comprehensive inpatient,

outpatient, or home-based PR programs, or it could be a training program targeting only LE exercise capacity. Again, the program had to include at least 20 regular sessions for a minimum frequency of 3 times a week; both supervised and unsupervised home sessions were acceptable. A detailed description of the control treatment was mandatory.

Outcome measures. These measures could be arm exercise capacity (maximal exercise capacity, functional exercise capacity, or endurance time). The UE *maximal exercise capacity* was defined as the peak exercise capacity measured by an incremental exercise stress test. *Functional exercise capacity* was defined as the maximum number of UE elevations performed in 6 minutes, and *arm endurance* was defined as the duration of a constant-load, symptom-limited exercise, performed using an arm ergometer. Outcome measures also could be the symptoms of dyspnea or arm fatigue on exertion, which had to be quantified by specific, validated questionnaires as scores achieved during exercises requiring exerting the UEs; the ability to perform ADL tasks that involve the arms, using reliable measures; or the HRQoL, as assessed by data collected by specific questionnaires.

Data Extraction and Quality Assessment

To assess eligibility, 2 investigators (SC and EC) independently retrieved and examined the titles and abstracts of the studies in order to achieve higher accuracy in this process; a third investigator (RD) was consulted in case of disagreement to improve accuracy. The 2 investigators extracted the data from the studies selected for inclusion and requested important data missing from these reports from their authors. Finally, the 2 investigators independently rated the quality of the studies selected using the Consolidated Standards of Reporting Trials (CONSORT) statement.[22]

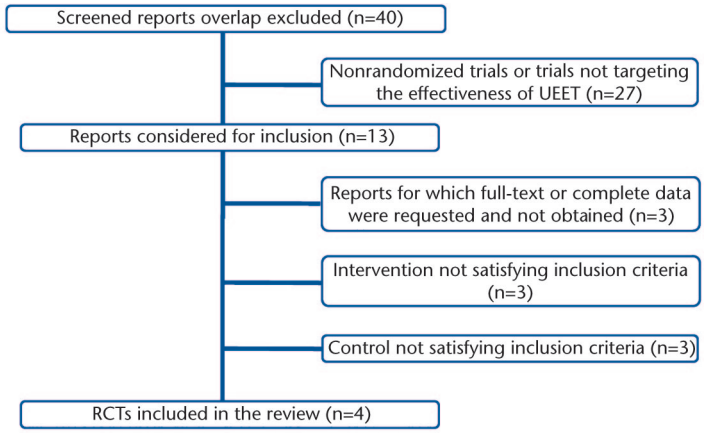

Figure.
Flow chart of screened, excluded, and eventually analyzed reports. Research articles included in more than one of the databases consulted (overlap) are not represented. RCT=randomized controlled trial, UEET=upper-extremity exercise training.

Results

Bibliographic Search Results

Forty publications were retrieved with combined computerized and hand search, including 2 abstracts and 1 reference (Figure). After a first review of the titles and abstracts, 13 studies remained potentially eligible. Efforts to obtain the full-length articles, or the data, directly from the authors of the reference and the 2 abstracts, presented at international meetings, were unsuccessful: in 2 cases, the authors did not answer our multiple attempts to contact them,[20,23] and in the third case, the authors were unable to provide the information requested.[24]

The 10 eligible full-text studies were independently reviewed as previously described. Agreement was reached to include 4 studies and to exclude 5 studies from the systematic review, using the Cohen coefficient of association (K=.8). One other study[25] was excluded after consulting the third investigator. Summarizing, 6 research articles were excluded because they did not satisfy inclusion criteria for intervention or control. Two trials did not focus on UEET as the experimental intervention,[25,26] and a third article was excluded because the intervention lasted less than 20 sessions,[27] thus not satisfying the evidence-based criteria for duration of PR as stated by the American Thoracic Society/European Respiratory Society[10] and because it included patients with diseases other than CAO. Two trials were excluded because the control groups did not perform comprehensive PR or LEET,[14,15] and the sixth trial was excluded because it compared 2 different modes of UEET, thus not allowing for control.[13]

One trial selected for this review[16] included 2 intervention groups, both satisfying our inclusion criteria, that were similar in the amount of training but different in the training modalities used. Both groups were separately considered for comparison with the control group. Another selected trial[12] included 4 groups for multiple comparisons. We checked them for inclusion criteria and considered the group that performed LEET only for comparison with the group that performed UEET plus LEET.

Table 1.

Characteristics of the Patient Samples and Study Designs of Trials That Fulfilled Eligibility Criteria[a]

Study	No. of Participants			Follow-up (wk)	Age (y), mean (SD)		Male/Female Ratio		Treatment	
	Randomized	At Completion	Allocated to Intervention Group		Control Group	Intervention Group	Control Group	Intervention Group	Control	Intervention
Ries et al, 1988[16]	45	28	8	8	Not reported		Not reported		PR (including LEET)	UEET (GRT) + GPT
			9							UEET (PNF) + GPT
Lake et al, 1990[12]	13	13	7	8	71.8 (3.3)	66.3 (6.8)	6/0	4/3	LEET	UEET + LEET
Sivori et al, 1998[28]	43	28	Not reported	8	63.0 (9.4)	66.07 (9.2)	12/2	11/3	LEET	UEET + LEET
Holland et al, 2004[17]	40	38	22	6	69.4 (6.6)	66.6 (8.4)	10/6	14/8	PR (including LEET) + placebo	UEET + LEET

[a] PR=comprehensive pulmonary rehabilitation, LEET=lower-extremity exercise training, UEET=upper-extremity exercise training, GRT=gravity-resistance training, GPT=general physical therapy, PNF=proprioceptive neuromuscular facilitation.

Table 2.

Characteristics of the Experimental Intervention in the Trials Included in the Review[a]

Study	No. of Sessions	Program Duration (wk)	Setting	Characteristics of Training
Ries et al, 1988[16]	55 (GRT)	6	Outpatient[b]	**Unsupported UEET (GRT):** Unsupported arm exercises in coordination with respiration to increase the endurance of the muscles of the arms and shoulders. Several repetitions against gravity only and against minimal resistance (1–5 lb). **Unsupported UEET (PNF):** Unsupported arm exercises in coordination with respiration against a progressive resistance on the basis of the PNF technique.
	28 (PNF)			
Lake et al, 1990[12]	24	8	Outpatient[b]	**Supported and unsupported UEET:** 10 min of warm-up followed by supported and unsupported arm exercises consisting of 20 min of circuit training (arm ergometer with variable resistance, ball throwing with horizontal arms, passing a beanbag overhead, tug of war linked to a pulley, dexterity exercises) followed by 10 min of cooling off.
Sivori et al, 1998[28]	24	8	Outpatient[b]	**Unsupported UEET:** A total of 20 min of training consisting of unsupported arm exercises in coordination with respiration. The sequence comprises the following 4 exercises: lifting a ball or a sandbag in front of the chest up to the head while maintaining the arms extended, lifting a wooden pole with the same technique, passing the ball from one hand to the other overhead, and passing the sandbag from one hand to the other with the same technique. Each exercise is repeated for 45 s, followed by 15 s of rest, until reaching 5 min of training. Increments are allowed on the basis of the patient's tolerance.
Holland et al, 2004[17]	30	6	Outpatient[b]/ home based	**Unsupported UEET:** Unsupported arm exercises aimed at increasing tolerance to effort and consisting of 3 min of weight lifting with an initial resistance of 500 g. The resistance is incremented by 0.5 kg in order to maintain the muscular effort between a grade of 12 to 14 and a dyspnea grade of at least 3 on the modified Borg self-administered scale.

[a] GRT=gravity-resistance training, PNF=proprioceptive neuromuscular facilitation, UEET=upper-extremity exercise training.
[b] Day hospital or outpatient visit.

Table 3.
Outcome Measures Used in the Trials Included in the Review[a]

Study	Upper-Extremity Exercise Capacity			Dyspnea on Exertion	Arm Fatigue on Exertion	ADL	HRQoL
	Maximal	Functional	Endurance Time				
Ries, 1988[16]	√		√	√ Borg-m	√ Borg-m	√	
Lake, 1990[12]	√			√	√ Borg		√ Bandura scale
Sivori, 1998[28]		√		√ Borg			√ CRDQ
Holland, 2004[17]	√			√ Borg-m	√ Borg		√ CRDQ

[a] ADL=activities of daily living, Borg=Borg dyspnea scale, Borg-m=modified Borg dyspnea scale (range=0–10), CRDQ=Chronic Respiratory Disease Questionnaire, HRQoL=health-related quality of life.

Characteristics of the Samples

Table 1 shows baseline demographic characteristics of participants in the studies, as well as study design features of the 4 RCTs that fulfilled all of the eligibility criteria. The sample size was small in all of the trials (13–45 participants); altogether, 141 participants were randomized, and 107 participants completed the trials (76% of those who had been randomized). Among the patients who completed the trials, 60 were allocated to the intervention group and 47 were allocated to the control group.

Participants were elderly and had severe or very severe CAO.[21] The main exclusion criteria in the selected trials were ischemic heart disease, heart failure, intermittent claudication, disabling musculoskeletal disorders, need for home oxygen treatment, hypercapnia, and medical conditions (other than CAO) severely limiting exercise tolerance.

Characteristics of the Training Programs

All of the selected trials included a detailed and complete description of the control and the intervention treatments (Tab. 2). Overall, UEET programs lasted 6 weeks[16,17] to 8 weeks[12,28] and included from 24 sessions[12,28] to 55 sessions.[16] In each study design, a short-term follow-up was planned, within 2 weeks from the end of the treatment.

Both the intervention and the control treatments were performed in an outpatient setting, although, in the study by Holland and co-workers,[17] an unsupervised, daily home exercise program integrated the twice-weekly, outpatient, supervised sessions. Upper-extremity training was conducted as unsupported arm exercises against gravity and progressive resistance as the major component in all of the trials. A combination of unsupported and supported UEET, using an arm ergometer, was used in one trial.[12]

Outcome Measures

Table 3 summarizes the outcome measures used to assess the treatment effects. Arm exercise capacity (maximal exercise tolerance, functional exercise tolerance, or arm endurance time) was measured in all of the trials. An incremental stress test, with either supported UEs[12,16] or unsupported UEs,[17] was performed in 3 trials to assess the maximal exercise capacity, whereas 1 trial[28] measured functional exercise capacity using a nonstandardized field test based on the number of arm elevations performed in 6 minutes. Finally, 1 trial[16] measured the duration of a constant load exercise, sustained by patients on the arm cycle, at a work level one step below their previously determined maximum. We did not consider the submaximal test performed on an arm ergometer in the trial by Lake and co-workers[12] because they did not report the endurance time.

Three trials[16,17,28] measured dyspnea on exertion, and 3 trials[12,16,17] measured arm fatigue on exertion, using different authorized versions of the Borg scale.[29,30] One trial[12] measured dyspnea with an unidentified scale; because the authors did not answer our request for clarification, the findings from this study are not reported for this specific outcome. The participants' ability to perform several ADL tasks, predominantly involving the arms, was measured in 1 trial only,[16] using a nonvalidated simulation field test. Three trials[12,17,28] evaluated HRQoL, using self-administered questionnaires.[31,32]

Methodological Quality of the Included Trials

Quality of reporting of the studies selected was assessed by means of the CONSORT statement,[22] which recently has been extended to cover reporting of nonpharmacological treatments such as physical therapy. It is a guideline designed to improve the reporting of RCTs. It consists of a checklist and a flow diagram that address the reporting of patient enrollment, allocation to treatments, follow-up, and data analysis. The CONSORT statement is widely used, and it is proven that the use of this evidence-based guideline is associated with improved quality of reporting of RCTs.[33]

Table 4.
Methodological Quality of the Trials Included in This Review[a]

Study	Section											
			Method									
	Title and Abstract	Introduction and Background	Participants	Interventions	Objectives	Outcomes	Sample Size	Randomization Sequence Generation	Randomization Allocation Concealment	Randomization Implementation	Blinding	Statistical Methods
Ries et al, 1988[16]	yes	YES	NO	yes	YES	no	NO	NO	NO	NO	NO	yes
Lake et al, 1990[12]	YES	YES	yes	yes	YES	no	NO	NO	NO	NO	NO	yes
Sivori et al, 1998[28]	yes	YES	yes	yes	YES	no	NO	yes	NO	NO	NO	yes
Holland et al, 2004[17]	yes	YES	yes	yes	YES	no	NO	NO	NO	NO	yes	yes

Table 4.
Continued

Study	Section										
	Results								Discussion		
	Participant Flow	Implementation of Intervention	Recruitment	Baseline Data	No. Analyzed	Outcomes and Estimation	Ancillary Analysis	Adverse Events	Interpretation	Generalizability	Overall Evidence
Ries et al, 1988[16]	no	YES	no	NO	no	YES	NA	YES	NO	yes	YES
Lake et al, 1990[12]	no	YES	no	yes	NO	YES	NA	YES	NO	yes	YES
Sivori et al, 1998[28]	no	NO	YES	yes	NO	YES	NA	YES	NO	NO	YES
Holland et al, 2004[17]	yes	NO	no	yes	yes	no	NA	YES	no	NO	YES

[a] NA=not applicable, no=criterion is satisfied for less than 50% of its components, NO=criterion is totally not satisfied, yes=criterion is satisfied for more than 50% of its components, YES=criterion is totally satisfied.

The selected trials completely satisfied a minimum of 4[17] and a maximum of 7[12] of the 23 evaluation criteria included (Tab. 4). In all cases, an accurate explanation of the rationale and hypothesis of the study was given, as well as precise details of the treatments provided and the statistical methods used. Furthermore, the results were always interpreted in the context of current evidence. However, a number of criteria were only partially met or not at all satisfied, introducing the possibility of systematic errors that reduce the internal validity of trials.[34] More specifically, in the trial of Ries and co-workers,[16] eligibility criteria for patients were not specified, clear definitions of primary and secondary outcome measures were not provided, and there was not a clear description of the flow of participants through each stage of the trial. Furthermore, their results were weakened by the high number of patients who dropped out and whose data were not collected at follow-up.

Similarly, the trials by Lake et al[12] and Sivori et al[28] did not provide clear definitions of primary and secondary outcome measures, did not describe the flow of participants through each stage of the trial, did not provide the number of participants included in each analysis for each group, and did not specify whether the statistical analysis followed an intention-to-treat approach. Also in the trial by Sivori and co-workers,[28] there was a large number of patients who dropped out and, therefore, were not reassessed at follow-up. Although Holland and co-workers[17] provided a clear description of eligibility criteria for their patients, clarified the primary and secondary outcome measures, provided a blinded assessment of the outcome measured, described the flow of participants through each phase of the trial, and followed an intention-to-treat approach, they did not provide an estimate of the effect size with its precision (eg, 95% confidence interval) for any of the outcomes measured.

As a whole, none of the selected trials specified how sample size was determined or the methods used to generate and implement the random allocation sequence. In 3 trials,[12,16,28] there was a complete lack of blinding for participants and the physical therapists administering the treatments and assessing the outcomes. One study[17] added a placebo treatment to the PR performed by the control group to disguise their allocation to this group. The placebo treatment consisted of finger dexterity exercises performed in a sitting position with arm supported and, therefore, was not expected to improve arm exercise capacity. Furthermore, we would like to point out that 3 trials[12,17,28] reported inequality in the male participant to female participant ratios, which might reduce the generalizability of their findings to the population of interest. The fourth trial[16] did not report these data.

Because of the poor methodological quality of the 4 RCTs included, we decided not to perform a meta-analysis, the results of which could be misleading given the low internal validity of the trials. However, we based the conclusions of our review on the results of the between-group comparisons made in each of the selected studies, and we always considered their limits.

Effects of UEET

The effects of UEET, compared with conventional PR or LEET, are summarized in Tables 5 and 6. Maximal exercise capacity was measured in 3 trials, and data were obtained from 79 participants with severe or very severe CAO. In 2 trials,[12,16] maximal exercise capacity was measured using an arm ergometer. In the third trial,[17] maximal exercise capacity was quantified as the duration of a standardized field test,[35] consisting of asking participants to raise their unsupported arms repeatedly, keeping an external pace, while the height of the target and the resistance were increased. Among the 3 trials mentioned, only 1 trial[17] detected a statistically significant increment of the maximal exercise capacity in favor of the intervention (change score=55.3 seconds, 95% confidence interval [CI]=8.25 to 102.35, $P<.02$). One trial[28] measured functional exercise capacity with a field test that satisfied our criteria. This outcome was collected in 28 individuals and documented a strong benefit in favor of the intervention group compared with the control group (change score=108, 95% CI=63.87 to 152.13, $P<.0001$), represented by an increased number of arm elevations per time unit. One trial[16] measured endurance time by registering the duration of a constant-load, symptom-limited exercise performed using an arm ergometer, and it did not show any statistically significant difference between groups.

Data regarding muscle effort of the UEs consistently showed no differences between intervention and control groups. Performance of ADL was measured in 28 patients with severe CAO enrolled in one trial.[16] This measurement was collected with a nonstandardized field test, simulating 3 common ADL tasks that involve the UEs and are usually considered to be fatiguing in this population. No statistically significant difference was detected between groups in this domain, either in terms of the time required to perform activities or as perceived symptoms.

With regard to symptoms during exertion, 3 trials[16,17,28] measured dyspnea and 3 trials[12,16,17] measured the effort of the UE muscles. Altogether, data regarding dyspnea were collected in 94 patients, and 79 patients

Table 5.
Effects of Upper-Extremity Exercise Training Plus Standard Training on Maximal and Functional Exercise Capacity, Endurance of the Upper Extremities, and Ability to Perform Activities of Daily Living (ADL) Involving the Upper Extremities[a]

Arm Exercise Capacity and Ability to Perform ADL						
Maximal exercise capacity: peak exercise capacity measured in watts,[16] Kpm/min,[12] or seconds[17] by incremental test						
Ries, 1988[16]	**Before**	**After**	**Mean Difference (CI) for Before-After Comparison**	***P* Value**	**Mean Difference (CI) for Between-Group Comparison**	***P* Value**
Control	16 (8)	20 (10)	4 (−3.57 to 11.57)	Not significant		Both comparisons not significant
Intervention-GR	16 (13)	17 (10)	1 (−10.37 to 12.37)	Not significant	−3.00 (−12.11 to 6.11)	
Intervention-PNF	13 (9)	12 (9)	−1 (−9.32 to 7.32)	Not significant	−8.00 (−16.34 to 0.34)	
Lake, 1990[12]	**Before**	**After**	**Mean Difference (CI) for Before-After Comparison**	***P* Value**	**Mean Difference (CI) for Between-Group Comparison**	***P* Value**
Control	27.0 (7.9)	27.1 (8.6)	0.10 (−9.24 to 9.44)	Not significant	3.20 (−6.75 to 13.15)	Not significant
Intervention	24.0 (11.1)	30.3 (9.7)	6.30 (−4.62 to 17.22)	*P*<.04		
Holland, 2004[17]	**Before**	**After**	**Mean Difference (CI) for Before-After Comparison**	***P* Value**	**Mean Difference (CI) for Between-Group Comparison**	***P* Value**
Control	Not reported	Not reported	Not estimated	Not reported	55.3 (8.25 to 102.35)	*P*<.02
Intervention	Not reported	Not reported	Not estimated	Not reported		
Functional exercise capacity: maximum number of upper-extremity elevations performed in 6 min						
Sivori, 1998[28]	**Before**	**After**	**Mean Difference (CI) for Before-After Comparison**	***P* Value**	**Mean Difference (CI) for Between-Group Comparison**	***P* Value**
Control	166.93 (58.38)	166.57 (58.24)	−0.36 (−39.78 to 39.06)	Not significant	108 (63.87 to 152.13)	*P*<.0001
Intervention	139.21 (45.64)	274.57 (60.86)	135.36 (98.21 to 172.5)	*P*<.0001		
Endurance: duration in seconds of a constant-load, symptom-limited exercise, performed using an arm ergometer						
Ries, 1988[16]	**Before**	**After**	**Mean Difference (CI) for Before-After Comparison**	***P* Value**	**Mean Difference (CI) for Between-Group Comparison**	***P* Value**
Control	185 (72)	181 (75)	−4 (−65.44 to 57.44)	Not significant		Both comparisons not significant
Intervention-GRT	215 (172)	195 (72)	−20 (−149.21 to 109.21)	Not significant	14 (−52.74 to 80.74)	
Intervention-PNF	135 (56)	144 (27)	9 (−31.62 to 49.62)	Not significant	−37 (−84.70 to 10.70)	
ADL: ability to perform ADL measured by the number of seconds needed to complete 3 tasks						
Ries, 1988[16]	**Before**	**After**	**Mean Difference (CI) for Before-After Comparison**	***P* Value**	**Mean Difference (CI) for Between-Group Comparison**	***P* Value**
Control	529 (86)	548 (96)	19 (−57.17 to 95.17)	Not significant		Both comparisons not significant
Intervention-GRT	665 (142)	663 (125)	−2 (−133.09 to 129.09)	Not significant	115 (11.45 to 218.55)	
Intervention-PNF	786 (361)	636 (234)	−150 (−431.07 to 131.07)	Not significant	88 (−75.07 to 251.07)	

[a] Results of comparisons within and between study groups for variables measured in the trials. Data are reported as mean (SD) or as mean (95% confidence interval [CI]). GRT=gravity-resistance training, PNF=proprioceptive neuromuscular facilitation, Kpm/min=kiloweight×meters/minute.

were assessed at follow-up for muscle effort. Although symptoms perceived during exertion always improved in both the intervention and control groups, a statistically significant difference in the dyspnea score, favoring the intervention group, was detected only in 1 trial[28] (change score=−1.07, 95% CI=−1.87 to −0.27, $P<.01$). Two trials detected no benefits in favor of either group.

Health-related quality of life was measured in 3 trials[12,17,28] for a total of 79 participants, using the Chronic Respiratory Disease Questionnaire[31] in 2 studies and the Bandura scale[32] in the third trial. Overall, the HRQoL improved for the intervention and control groups in all studies at the end of the treatment, but none of the studies

Table 6.
Effects of Upper-Extremity Exercise Training Plus Standard Training on the Symptoms of Dyspnea and Arm Fatigue During Activities Involving the Upper Extremities and on Health-Related Quality of Life (HRQoL)[a]

Symptoms and HRQoL						
Dyspnea: score achieved during exercise exerting the upper extremities, as measured by the Borg scale[28] or the modified Borg scale[16,17]						
Ries et al, 1988[16]	**Before**	**After**	**Mean Difference (CI) for Before-After Comparison**	***P* Value**	**Mean Difference (CI) for Between-Group Comparison**	***P* value**
Control	5.5 (2.9)	4.1 (1.8)	−1.40 (−3.42 to 0.62)	*P*<.05		Both comparisons not significant
Intervention-GRT	4.9 (2.0)	3.3 (0.8)	−1.60 (−3.09 to −0.11)	*P*<.05	−0.80 (−2.0 to 0.4)	
Intervention-PNF	4.9 (2.0)	4.1 (1.4)	−0.80 (−2.39 to 0.79)	*P*<.05	0.00 (−1.4 to 1.4)	
Holland et al, 2004[17]	**Δ Before-After**		**Mean Difference (CI) for Before-After Comparison**	***P* Value**	**Mean Difference (CI) for Between-Group Comparison**	***P* Value**
Control	Δ −2.9 (0.78)		−2.90 (−3.28 to −2.52)	Not reported	Not estimated	Not significant
Intervention	Δ −4.0 (0.84)		−4.00 (−11.72 to 3.72)	Not reported		
Sivori et al, 1998[28]	**Before**	**After**	**Mean Difference (CI) for Before-After Comparison**	***P* Value**	**Mean Difference (CI) for Between-Group Comparison**	***P* value**
Control	2.21 (1.76)	1.86 (1.17)	−0.35 (−1.32 to 0.62)	Not significant	−1.07 (−1.87 to −0.27)	*P*<.01
Intervention	2.50 (1.79)	0.79 (0.97)	−1.71 (−2.61 to −0.81)	*P*<.001		
Arm fatigue: score achieved during exercise exerting the upper extremities, as measured by the Borg scale[12,17] or the modified Borg scale[16]						
Ries et al, 1988[16]	**Before**	**After**	**Mean Difference (CI) for Before-After Comparison**	***P* Value**	**Mean Difference (CI) for Between-Group Comparison**	***P* Value**
Control	5.5 (2.4)	5.1 (2.2)	−0.40 (−2.32 to 1.52)	Not significant		Both comparisons not significant
Intervention-GRT	4.6 (2.4)	4.2 (2.1)	−0.40 (−2.61 to 1.81)	Not significant	−0.90 (−2.85 to 1.05)	
Intervention-PNF	4.3 (1.8)	3.6 (1.3)	−0.70 (−2.15 to 0.75)	Not significant	−1.50 (−3.05 to 0.05)	
Lake et al, 1990[12]	**Before**	**After**	**Mean Difference (CI) for Before-After Comparison**	***P* Value**	**Mean Difference (CI) for Between-Group Comparison**	***P* Value**
Control	12.7 (0.5)	13.2 (1.1)	0.50 (−0.47 to 1.47)	Not significant	−1.20 (−2.56 to 0.16)	Not significant
Intervention	12.1 (0.8)	12.0 (1.4)	−0.10 (−1.29 to 1.09)	Not significant		
Holland et al, 2004[17]	**Before**	**After**	**Mean Difference (CI) for Before-After Comparison**	***P* Value**	**Mean Difference (CI) for Between-Group Comparison**	***P* Value**
Control	Not reported	Not reported	Not estimated	Not reported	Not estimated	Not significant*
Intervention	Not reported	Not reported	Not estimated	Not reported		

(continued)

revealed a statistically significant improvement in the intervention group compared with the control group.

Discussion and Conclusions

This systematic review demonstrates that there is insufficient evidence to support the inclusion of UEET in PR programs for patients with severe and very severe CAO. Although the results of 2 trials included,[17,28] when considered separately, may suggest some advantages when UEET is incorporated into standard PR programs, the same results, when taken together, are strongly contradictory and, therefore, inadequate to recommend this activity.

Due to numerous shortcomings existing in the 4 RCTs included, the overall quality of the evidence collected in this systematic review was low for any of the outcomes studied. By examining the influence of key components of study quality for each trial reviewed, we found that potential sources of selection bias might exist in all of the selected trials. Se-

Table 6.
Continued

HRQoL: score achieved by specific, validated questionnaires						
Lake et al, 1990[12]	**Δ Before-After**		**Mean Difference (CI) for Before-After Comparison**	***P* Value**	**Mean Difference (CI) for Between-Group Comparison**	***P* Value**
Control	Δ +7%		Not estimated	Not reported	Not estimated	Not significant*
Intervention	Δ +24%		Not estimated	*P*<.005*		
Holland et al, 2004[17]	**Before**	**After**	**Mean Difference (CI) for Before-After Comparison**	***P* Value**	**Mean Difference (CI) for Between-Group Comparison**	***P* Value**
Control	Not reported	Not reported	Not estimated	Not reported	Not estimated	Not significant*
Intervention	Not reported	Not reported	Not estimated	Not reported		
Sivori et al, 1998[28]	**Before**	**After**	**Mean Difference (CI) for Before-After Comparison**	***P* Value**	**Mean Difference (CI) for Between-Group Comparison**	***P* Value**
Control	87.57 (29.81)	111.79 (17.29)	24.22 (8.58 to 39.86)	*P*<.0001	−4.29 (−16.98 to 8.40)	Not significant
Intervention	75.14 (24.74)	107.5 (16.96)	32.3 (18.73 to 45.99)	*P*<.0001		

[a] Results of comparisons within and between study groups for variables measured in the trials. Data are reported as mean (SD) or as mean (95% confidence interval [CI]), GRT=gravity-resistance training, PNF=proprioceptive neuromuscular facilitation, Δ=difference between means. Asterisk indicates as reported by the authors in the text.

quence of allocation to treatments was not concealed in any of the RCTs, although it is well known that investigators' knowledge of the sequence of allocation may cause selective enrollment of patients on the basis of prognostic factors[34] and, consequently, may lead to inflated treatment effects.[34]

Performance bias and detection bias arise when the lack of double-blinding influences additional treatments that might be offered preferentially to one group or the assessment of the outcomes, respectively. Again, such biases may inflate treatment effects to a different degree, depending on the outcome assessed. This possibility is strongly reduced by the blinding of those administering the treatment, which is almost impossible in the field of physical therapy, and the blinding of patients and those assessing outcomes, which was accomplished by 1 trial[17] among the 4 trials selected. The same trial was the only one that minimized the sources of attrition bias by making every effort to reduce the number of data lost to follow-up (1%) and using an intention-to-treat approach. Conversely, reporting of 3 trials[12,16,28] was unclear regarding the approach followed in the data analysis, and in 2 trials[16,28] the participant dropout rate was higher than 30%, thus reducing the validity of the findings. Moreover, not even one of the studies selected stated the intended sample size, and some trials needed multiple comparisons because they included more than one intervention group[16,12] or control group,[12] thus reducing the power of the analysis.

If UEET did indeed result in improvements, these changes may not have been able to be identified for several reasons, including the lack of sample power in all of these trials. Additionally, all of the trials implemented UEET specifically targeted at submaximal performance levels, whereas the testing procedures measured maximal performance. Finally, the benefit of UEET in the experimental group on the outcomes of HRQoL, dyspnea, and arm fatigue may have been masked by the concomitant participation in a PR program, which is known to improve these measures and would have done so in both experimental and control groups.

Thus, the overall quality of the trials included in this review is very low in 3 cases[12,16,28] and unsatisfactory in the fourth case.[17] Taking all of these weaknesses into consideration, the findings of this review cannot support the inclusion of UEET in PR programs for patients with CAO, and even the inconsistent advantages shown individually by some of the trials included[12,17,28] may be overestimated.[34]

Notwithstanding these findings, the most recent guidelines for PR[10,11] strongly recommend the introduction of unsupported UEET of sufficient duration (ie, 20 sessions) in PR programs. This was the main reason why we decided to include trials with at least 20 sessions of UEET in our review analysis, thus excluding other studies of different duration. Despite the belief that the longer the program, the greater the benefit,[10] we cannot deny that significant benefits of UEET may occur in trials of

shorter duration, such as in the study by Porta et al[27]; however, that particular study was performed in a very different population.

The rationale that supports the inclusion of specific training directed at the UEs in patients with CAO in PR programs relies on data from 6 randomized studies[12–17] (3 included in this review) and 3 nonrandomized studies.[18–20] The underlying principle is that an improvement in arm exercise capacity might be particularly important in these patients, whose UE muscles are competitively involved in both arm elevation and accessory ventilation. The same guidelines[11] postulate that the mechanisms for improvement in UE function from such training include desensitization to dyspnea, better muscular coordination, and metabolic adaptation to exercise.

Numerous patients with stable, moderate to severe CAO complain of dyspnea during activities involving the UEs. These patients often show a characteristic association of dyspnea, dyssynchronous breathing,[1] and inefficient metabolic and ventilatory response.[1,36,37] A possible explanation of these phenomena is that, in patients with CAO who have hyperinflation, the diaphragm is less effective in performing inspiration.[6] Consequently, during unsupported arm activities, these patients, compared with people who are healthy, must rely more on the accessory inspiratory muscles, which are involved in the competitive demands of ventilation and arm elevation. This fact poses greater demands on the accessory inspiratory muscle function, thus sustaining the hypothesis that these multifunctional muscles might benefit from specific training.[38] However, the symptom of dyspnea was assessed in 3 trials[16,17,28] based on unsupported UEET, and a statistically and clinically[39] significant benefit in favor of the intervention was detected in only 1 trial.[28] Taken together, the results reported in this review cannot support or refute the hypothesis that arm exercise may improve dyspnea.

The ultimate scope of rehabilitation is to improve the patient's autonomy in daily life. Pulmonary rehabilitation and exercise training, in particular, contribute in a decisive way to this process. Any accomplishment in this domain should be demonstrated by an increase in the patient's ability to perform ADL in his or her own environment. Unfortunately, the RCTs included in this review did not investigate these areas, and the trial[16] that assessed the ability to perform ADL with a nonstandardized field test was unable to detect any favorable effects of UEET. Furthermore, arm fatigue was unchanged by the addition of UEET to a standard PR program. Interestingly, one trial screened and excluded from this review[15] showed that, when UEET was administered independent of standard PR, it failed to provide any benefit compared with a control treatment. This finding suggests that UEET alone is not sufficient to improve clinically important outcomes for patients with CAO.

However, when implemented in the unsupported modality, UEET may add additional benefit to the established results of standard PR programs in terms of maximal and functional exercise capacity of the UEs. Indeed, Sivori and colleagues[28] demonstrated a 100% improvement in functional exercise capacity of the arm, which was associated with a decrease in dyspnea but no change in HRQoL. Similarly, when a higher peak of exercise capacity was documented due to the effect of UEET, it did not translate to a reduction of dyspnea or arm fatigue, nor did it lead to a better HRQoL. Therefore, the effect of UEET on maximal exercise capacity is equivocal; findings in favor of UEET detected by individual trials are difficult to interpret and would exclude that desensitization to dyspnea and metabolic adaptation to exercise are possible mechanisms of improvement in UE exercise capacity, as recently postulated.[11]

This review did not demonstrate any additional improvement in HRQoL for patients who underwent UEET. In fact, among 3 trials[12,17,28] that measured HRQoL, none showed a significant difference in favor of the intervention group. However, the trial by Lake and colleagues[12] showed a trend favoring the intervention group in comparison with the control group (24% versus 7%). This benefit could be due to the Hawthorne effect, because in this trial patients were not blinded to treatment group allocation and no other improvement was detected in the other outcomes measured to substantiate this finding.

To our knowledge, this is the first systematic review examining the effectiveness of UEET in patients with CAO that has been performed using a rigorous, yet broad, search in different languages. The available evidence is limited, and the outcome measures examined varied considerably. Furthermore, the possibility that the samples were heterogeneous, coupled with the diverse UEET training protocols, limits the aggregation of the data. Finally, the relatively poor methodological quality of the included studies compromised both internal validity and generalizability of the results. These factors prevented us from conducting a meta-analysis, which might have been useful in clarifying the efficacy of UEET in patients with CAO.

In summary, the available evidence from RCTs appears inadequate to recommend in favor of or against the inclusion of UEET in PR programs for individuals with CAO. Further research should be conducted by

means of well-designed and adequately powered trials, based on validated outcome measures addressed to clinically meaningful end points. The development of standardized and quantitative tests to assess the ability of patients with CAO to perform ADL also would be helpful to obtain a deeper understanding of clinically important achievements from the patients' perspectives. Other important related research questions should be whether patients with different CAO severity or levels of disability might benefit differently from UEET and whether unsupported versus supported arm exercise might provide greater, or more selective, benefits.

Ms Costi, Dr Di Bari, Dr Crisafulli, Dr Fabbri, and Dr Clini provided concept/idea/research design. Ms Costi, Dr Di Bari, Mr Pillastrini, and Dr Clini provided writing. Ms Costi, Dr Di Bari, Dr Crisafulli, and Ms Arletti provided data collection. Ms Costi, Dr Di Bari, Dr D'Amico, and Ms Arletti provided data analysis. Ms Costi, Dr Di Bari, Mr Pillastrini, and Dr Crisafulli provided project management. Ms Costi, Dr Fabbri, and Dr Clini provided fund procurement, facilities/equipment, and institutional liaisons. Ms Costi, Dr D'Amico, and Dr Clini provided participants. The authors thank Mrs Jessie Cross for her helpful editorial advice and linguistic review of the manuscript.

This article was received December 16, 2007, and was accepted January 29, 2009.

DOI: 10.2522/ptj.20070368

References

1 Celli BR, Rassulo J, Make BJ. Dyssynchronous breathing during arm but not leg exercise in patients with chronic airflow obstruction. *N Engl J Med*. 1986;314:1485–1490.

2 Criner GJ, Celli BR. Effect of unsupported arm exercise on ventilatory muscle recruitment in patients with severe chronic airflow obstruction. *Am Rev Respir Dis.* 1988;138:856–861.

3 Martinez FJ, Strawderman RL, Flaherty KR, et al. Respiratory response during arm elevation in isolated diaphragm weakness. *Am J Respir Crit Care Med.* 1999;160:480–486.

4 Sharp JT, Danon J, Druz WS, et al. Respiratory muscle function in patients with chronic obstructive pulmonary disease: its relationship to disability and to respiratory therapy. *Am Rev Respir Dis.* 1974;110:154–167.

5 Celli BR. The clinical use of upper extremity exercise. *Clin Chest Med.* 1994;15:339–349.

6 Polkey MI, Kyroussis D, Hamnegard CH, et al. Diaphragm strength in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med.* 1996;154:1310–1317.

7 Hamilton AL, Killian KJ, Summers E, et al. Muscle strength, symptom intensity and exercise capacity in patients with cardiorespiratory disorders. *Am J Respir Crit Care Med.* 1995;152:2021–2031

8 O'Donnell DE, Bertley JC, Chau LK, et al. Qualitative aspects of exertional breathlessness in chronic airflow limitation: pathophysiologic mechanisms. *Am J Respir Crit Care Med.* 1997;155:109–115.

9 Lacasse Y, Brosseau L, Milne S, et al. Pulmonary rehabilitation for chronic obstructive pulmonary disease. *Cochrane Database Syst Rev.* 2002;3:CD003793.

10 Nici L, Donner C, Wouters E, et al. American Thoracic Society/European Respiratory Society statement on pulmonary rehabilitation. *Am J Respir Crit Care Med.* 2006;173:1390–413.

11 Ries AL, Bauldoff GS, Carlin BW, et al. Pulmonary Rehabilitation: Joint ACCP/AACVPR Evidence-Based Clinical Practice Guidelines. *Chest.* 2007;131:4S–42S

12 Lake F, Henderson K, Briffa T, et al. Upper limb and lower limb exercise training in patients with chronic airflow obstruction. *Chest.* 1990;97:1077–1082.

13 Martinez FJ, Vogel PD, Dupont DN, et al. Supported arm exercise vs unsupported arm exercise in the rehabilitation of patients with severe chronic airflow obstruction. *Chest.* 1993;103:1397–1402.

14 Epstein S, Celli BR, Martinez FJ, et al. Arm training reduces the VO2 and VE cost of unsupported arm exercise and elevation in chronic obstructive pulmonary disease. *J Cardiopulmonary Rehabil.* 1997;17:171–177.

15 Bauldoff GS, Hoffman L, Sciurba F, et al. Home-based, upper arm exercise training for patients with chronic obstructive pulmonary disease. *Heart Lung.* 1996;25:288–294.

16 Ries AL, Ellis B, Hawkins RW. Upper extremity exercise training in chronic obstructive pulmonary disease. *Chest.* 1988;93:688–692.

17 Holland AE, Hill CJ, Nehez E, et al. Does unsupported upper limb exercise training improve symptoms and quality of life for patients with chronic obstructive pulmonary disease? *J Cardiopulmonary Rehabil*. 2004;24:422–427.

18 Franssen FM, Wouters EF, Baarends EM, et al. Arm mechanical efficiency and arm exercise capacity are relatively preserved in chronic obstructive pulmonary disease. *Med Sci Sports Exerc.* 2002;34:1570–1576.

19 Couser JI Jr, Martinez FJ, Celli BR. Pulmonary rehabilitation that includes arm exercise reduces metabolic and ventilatory requirements for simple arm elevation. *Chest.* 1993;103:37–41.

20 O'Hara WJ, Lasachuk KE, Matheson PC, et al. Weight training and backpacking in COPD. *Respir Care.* 1984;29:1202–1210.

21 Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: NHLBI/WHO Global Initiative for Chronic Obstructive Lung Disease (GOLD) Workshop Summary. *Am J Respir Crit Care Med.* 2001;163:1256–1276.

22 Boutron I, Moher D, Altman DG, Schulz KF, Ravaud P; CONSORT Group. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Ann Intern Med*. 2008;148:295–309.

23 Shu MF, Kao CH, Kuo HP. Upper arm exercise improves exercise tolerance and dyspnea sensation in patients with chronic obstructive airway disease (COAD) [abstract]. *Eur Respir J.* 1998;12(suppl):406S.

24 Baarends BM, Crautzherg EC, Janssen PP, et al. Functional effects of unsupported arm exercise training in addition to nutritional therapy in depleted patients with chronic obstructive pulmonary disease (COPD) participating in a pulmonary rehabilitation program [abstract]. *Eur Respir J.* 1999;13(suppl):211S.

25 Normandin EA, McCusker C, Connors M, et al. An evaluation of two approaches to exercise conditioning in pulmonary rehabilitation. *Chest.* 2002;121:1085–1091.

26 Bauldoff G, Rittinger M, Nelson T, et al. Feasibility of distractive auditory stimuli on upper extremity training in persons with chronic obstructive pulmonary disease. *J Cardiopulmonary Rehabil*. 2005;25:50–55.

27 Porta R, Vitacca M, Gilè LS, et al. Supported arm training in patients recently weaned from mechanical ventilation. *Chest.* 2005;128:2511–2520.

28 Sivori M, Rhodius E, Kaplan P, et al. Exercise training in chronic obstructive pulmonary disease. *Medicina (B Aires).* 1998;58:717–727.

29 Borg GA. Perceived exertion. *Exerc Sport Sci Rev.* 1974;2:131–153.

30 Borg GA. Psychophysical bases of perceived exertion. *Med Sci Sports Exerc.* 1982;14;377–381.

31 Guyatt GH, Berman LB, Townsend M, et al. A measure of quality of life for clinical trials in chronic lung disease. *Thorax.* 1987;42:773–778.

32 Bandura A, Adam NE. Analysis of self-efficacy theory of behavioural change. *Cognit Ther Res.* 1977;1:287–310.

33 Plint AC, Moher D, Morrison A, et al. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. *Med J Aust*. 2006;185:263–267.

34 Juni P, Altman DG, Egger M. Assessing the quality of controlled clinical trials. *BMJ*. 2001;323:42–46.

35 Takahashi T, Jenkins SC, Strauss GR, et al. A new unsupported upper limb exercise test for patients with chronic obstructive pulmonary disease. *J Cardiopulm Rehabil*. 2003;23:430–437.

36 Couser JI, Martinez FJ, Celli BR. Respiratory response and ventilatory muscle recruitment during arm elevation in normal subjects. *Chest*. 1992;101:336–340.

37 Dolmage TE, Avendano MA, Goldstein RS. The ventilatory response to arm elevation of patients with chronic obstructive pulmonary disease. *Chest*. 1993;1047:1097–1100.

38 Gigliotti F, Coli C, Bianchi R, et al. Arm exercise and hyperinflation in patients with COPD: effect of arm training. *Chest*. 2005;128:1225–1232.

39 Ries AL. Minimally clinically important difference for the UCSD Shortness of Breath Questionnaire, Borg Scale, and Visual Analog Scale. *COPD*. 2005;2:105–110.